Big Mart Sales Prediction

In the previous module we solved the Loan Prediction problem. That was a classification problem where we had to predict whether the loan should be approved or not.

Similar to that problem, here we have an assignment for you where you have to predict the sales of a big mart.

Problem Statement

The data scientists at BigMart have collected 2013 sales data for 1559 products across 10 stores in different cities. Also, certain attributes of each product and store have been defined. The aim is to build a predictive model and find out the sales of each product at a particular store.

About Dataset

We have train and test data set. Train data set has both input and output variable. You need to predict the sales for test data set.

Number of observations in Train set: 8523

Number of observations in Test set: 5681

Data Download

You can download the dataset from our [Datahack Platform](). Several Machine Learning and Deep Learning competitions are hosted on this platform. It gives you the following functionalities:

- You can view and compare your score against that of other participants on the leaderboard.
- You can also keep track of your submissions.

Approach to the Problem

1. Pre-process the data: impute the missing values, normalize the variables. etc.
2. Define the architecture of your model
3. Since it is a regression problem, make sure that you use the 'linear' activation function in the output layer
4. Train your model
5. Do the same pre-processing steps on test set as you did for the training set
6. Generate predictions for the test set using the trained model
7. Save the predictions in a csv file (to check the format, refer to the sample submission file provided on the problem page)
8. Submit your predictions on the problem page and check your rank on the leaderboard.